



# An Open Source web-based collection management system for natural history specimens and for publishing data at GBIF

Ingimar Erlingsson / Ингимар Эрлингссон

Systems Developer in the DINA-team  
Bioinformatics and Genetics at the Swedish Museum of Natural History

[Ingimar.Erlingsson@nrm.se](mailto:Ingimar.Erlingsson@nrm.se)

# Introduction

Presentation of The Swedish Museum of Natural History  
Presentation of The DINA Collection Management System





# The Swedish Museum of Natural History

Staff : 260 persons, 6 research departments  
2/3 are working in the research departments

- The units at the museum
  - Bioinformatics and Genetics, Staff : 30 persons
  - Botany
  - Geosciences
  - Zoology
  - Palaeobiology
  - Environmental research and monitoring

# Collections at NRM: Total of 9 767 700 specimens\*

<b>Botany</b>	
- Flowering plants and Ferns	3,000,000
- Mosses, Fungi, Algae	1,500,000
<b>Zoology</b>	
- Vertebrates	300,000
- Evertebrates	550,000
- Entomology	2,500,000
<b>Paleo</b>	
- PaleoZoology	860,000
- PaleoBotany	175,000
<b>Geology</b>	150,000
Environmental Specimen Bank	~ 270,000 accessions

Approx. 33 Million specimens in  
total from Sweden  
Statistics from 2013

# Bioinformatics and Genetics (BIO) , 30 persons

## Activity in our department ....

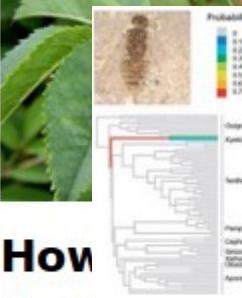


### Automatic Species Identification

We teach robots to identify insects, so that experts with round identification keys can focus on more difficult identifications.

#### How was it done?

**Phylogenetic analysis** is used to understand evolutionary relationships between species. The software involves Bayesian phylogenetic management system. This is widely used in medicine and biology.



### DINA web-based Collections Management System

Leading international cooperation on the development of a new Open Source collections management system.

The software involves Bayesian phylogenetic management system. This is widely used in medicine and biology.



### BioAtlas Sweden / GBIF-Sweden

Biodiversity Atlas Sweden (BAS including GBIF-Sweden) provides data and analysis services for Swedish and international biodiversity and species in ecosystems. Our researches are significant in previously unexplored areas.

### Insects in Sweden

Our researches are significant in previously unexplored areas.

cesses affect red species?



# The DINA-Mission

## The current situation [Sweden]

- Many Natural History Museums
- Stand-alone systems, '*home grown*' collection management system

### Problems

- People-centric
- Cost efficient ?
  - maintenance & operations
  - user support
- Vendor lock-in, license-cost
- Challenges when sharing data

## Our Vision

- All NHM:s will use the same system
- A web-based collection management system

### Solutions

- Organization-centric
- Cost efficient!
  - Centralized maintenance & operations
  - Centralized user support
- Open Source, no license-fee
- Ease of Sharing Data (e.g to GBIF)





## The DINA-Vision

We are building a **platform** ...

A web-based collection management system

...for all collections...

...in the long term, replacing one system at a time

A more efficient collections management

→ No need for managing multiple systems

To open up the data to the research community and to the public!



# August 2017 , Restart of the Project

- Always been Open Source
  - Code is available at GitHub and binaries at Docker Hub
- Agile way
  - Sprint / Sprint-planning / daily standup / Sprint-review with demo / retrospect / a digital 'Kanban board' ....
- Change of Technology Stack , All JavaScript
- Team and the Roles.
  - 1 Project manager, 2 domain specialists , 4 developers
  - All at the same location
- MVP (Minimal Viable Product)
  - Defining the MVP
  - Concentrating on a single collection



# A full JavaScript Stack

- **Frontend core technologies**
    - React & Redux
    - Semantic UI
    - JSON schema
  - **Backend core technologies**
    - NodeJS
    - Express
    - JSON schema (standard)
    - Data store
  - **Devops**
    - Linux Servers
    - Docker containers
      - "DINA"-cm, nginx, Keycloak
  - **Data storage**
    - Elasticsearch
    - Postgres
- Rest Conventions → JSON API*



# Short DEMO 1 (2)

Specimens Mammals

24662 Total records [Show All](#) [+ New record](#) [Q.Find](#)

Row #	Catalog number	Taxon (curatorial name)	Family	Genus	Species	Subspecies	Collectors	Type status
1	590774	Galerella sanguinea	Herpestidae	Galerella	Galerella sanguinea		HRH Prince Wilhelm	
2	590772	Myotis daubentonii	Vespertilionidae	Myotis	Myotis daubentonii			
3	558400	Physeter catodon	Physeteridae	Physeter	Physeter catodon			
4	590765						Söderström, L	
5	590764						Söderström, L	
6	590762	Nectomys squamipes	Cricetidae	Nectomys	Nectomys squami...		von Ihering, Hermann	
7	558399	Balaenoptera edeni	Balaenopteridae	Balaenoptera	Balaenoptera edeni		Johnson, CO	
8	590755	Potamochoerus porcus	Suidae	Potamochoerus	Potamochoerus po...		Waldau, Georg	
9	590752	Martes martes	Mustelidae	Martes	Martes martes			
10	590751	Pteromys volans	Sciuridae	Pteromys	Pteromys volans			
11	590750	Chironectes minimus	Didelphidae	Chironectes	Chironectes minim...		Bovallius, Carl	
12	590745						Lönnberg, Einar	
13	590741						Bergman, Sten	



# Short DEMO 2 (2)

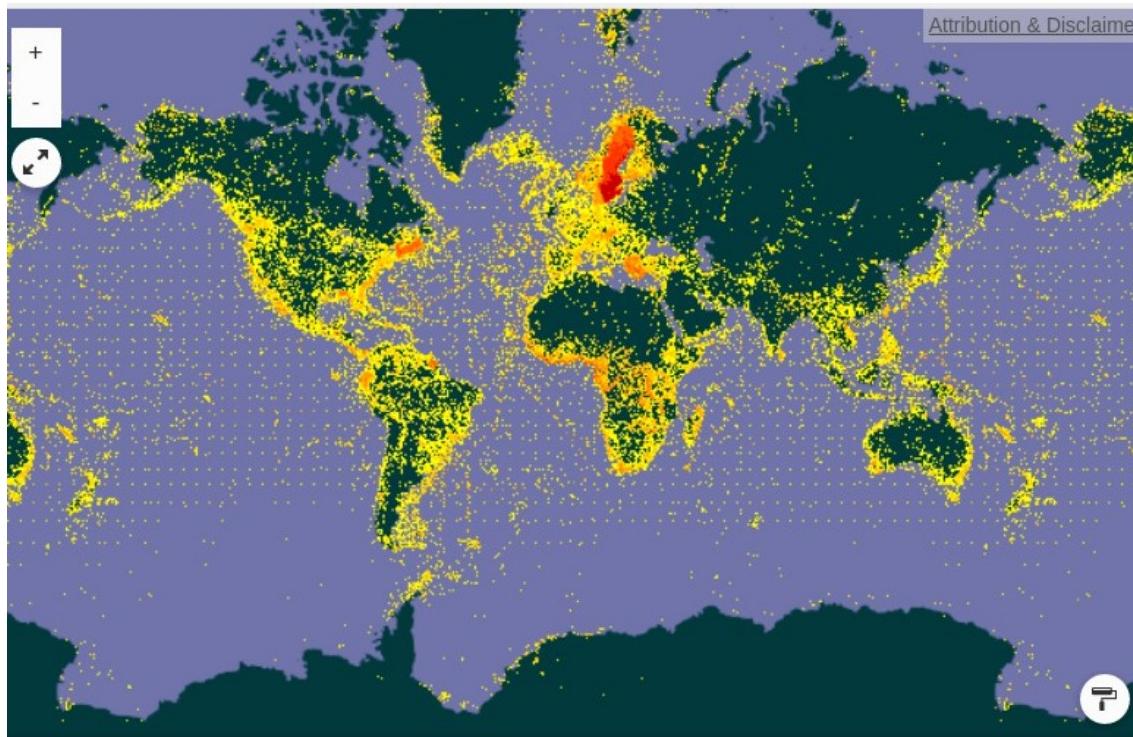
Specimens Mammals

24662 Total records [Show All](#) [+ New record](#) [Q Find](#)

Row #	Catalog number	Taxon (curatorial name)	Family	Genus	Species	Subspecies	Collectors	Type status	Start date (collecting)	End date (collecting)	Locality (collecting)	Count (collector)
1	590774	Galerella sanguinea	Herpestidae	Galerella	Galerella sanguinea		HRH Prince Wilhelm		1914-02-07			Kenya
2	590772	Myotis daubentonii	Vespertilionidae	Myotis	Myotis daubentonii				1830-10-26		Mörkö	Sweden
3	558400	Physeter catodon	Physeteridae	Physeter	Physeter catodon				1905-01-01		Färöarna	Denmark
4	590765					Söderström, L.			1912-05-10		Ecuador; Pichincha	Ecuador
5	590764					Söderström, L.			1913-04-10		Ecuador; Pichincha ?	Ecuador
6	590762	Nectomys squamipes	Cricetidae	Nectomys	Nectomys squami...		von Ihering, Hermann				Taquara do Mundo novo	Brazil
7	558399	Balaenoptera edeni	Balaenopteridae	Balaenoptera	Balaenoptera edeni		Johnson, CO		1929-03-10		Sydafrika	South Africa
8	590755	Potamochoerus porcus	Suidae	Potamochoerus	Potamochoerus po...		Waldau, Georg		1840-01-01			South Africa
9	590752	Martes martes	Mustelidae	Martes	Martes martes				1838-01-01		Jämtland	Sweden
10	590751	Pteromys volans	Sciuridae	Pteromys	Pteromys volans						Finland, Kuopio	Finland
11	590750	Chironectes minimus	Didelphidae	Chironectes	Chironectes minim...		Bovallius, Carl				Britiska Guyana	Guyana
12	590745					Lönnberg, Einar			1911-02-08		Lekkundo River, Kenya Col.	Kenya
13	590741					Bergman, Sten			1936-06-16		Ringam po, Korea	Korea
14	558398	Mesoplodon bidens	Ziphidae	Mesoplodon	Mesoplodon bidens		Aurivillius		1885-08-06			Sweden
15	590732	Mungos mungo	Herpestidae	Mungos	Mungos mungo		Wahlberg, Johan August		1839-08-23		Durban//	South Africa
16	590731	Mustela putorius	Mustelidae	Mustela	Mustela putorius		Swedish Museum of Natural Histor...		1914-03-27		Bara/Torup//	Sweden
17	590729	Galerella sanguinea	Herpestidae	Galerella	Galerella sanguinea		Lönnberg, Einar		1911-03-11			Kenya
18	590721					Högfeldt, P.J.			1926-10-06		Nyborg, Hätuna	Sweden
19	590720	Procavia capensis	Procaviidae	Procavia	Procavia capensis		Wahlberg, Johan August		1835-04-01		Sydafrika kapkolonien, Kalk bay	South Africa
20	558397	Balaenoptera acutorostrata	Balaenopteridae	Balaenoptera	Balaenoptera acut...		Collector, anonymous		1879-01-01			Sweden
21	590719	Procavia capensis	Procaviidae	Procavia	Procavia capensis		Wahlberg, Johan August		1835-04-01			South Africa
22	590718					Beck-Friis, C.			1907-01-05		Böringe	Sweden
23	590717					Thott, Tage			1908-10-27		Skabersjö	Sweden



# GBIF-Sweden, 2017



## Data from Sweden

- 43 occurrence datasets with 61,523,442 records.
- 1 checklist dataset with 105,768 records.
- No metadata-only datasets.
- Sweden publishes data covering 243 countries, territories and islands.

[View records shown on the map](#)

Source : <http://www.gbif.org/resource/81771>

Source : Sweden → <http://www.gbif.org/country/SE/summary> , <https://demo.gbif.org/country/SE>



# GBIF Sweden (2018)



81,356,416

Published occurrences

44

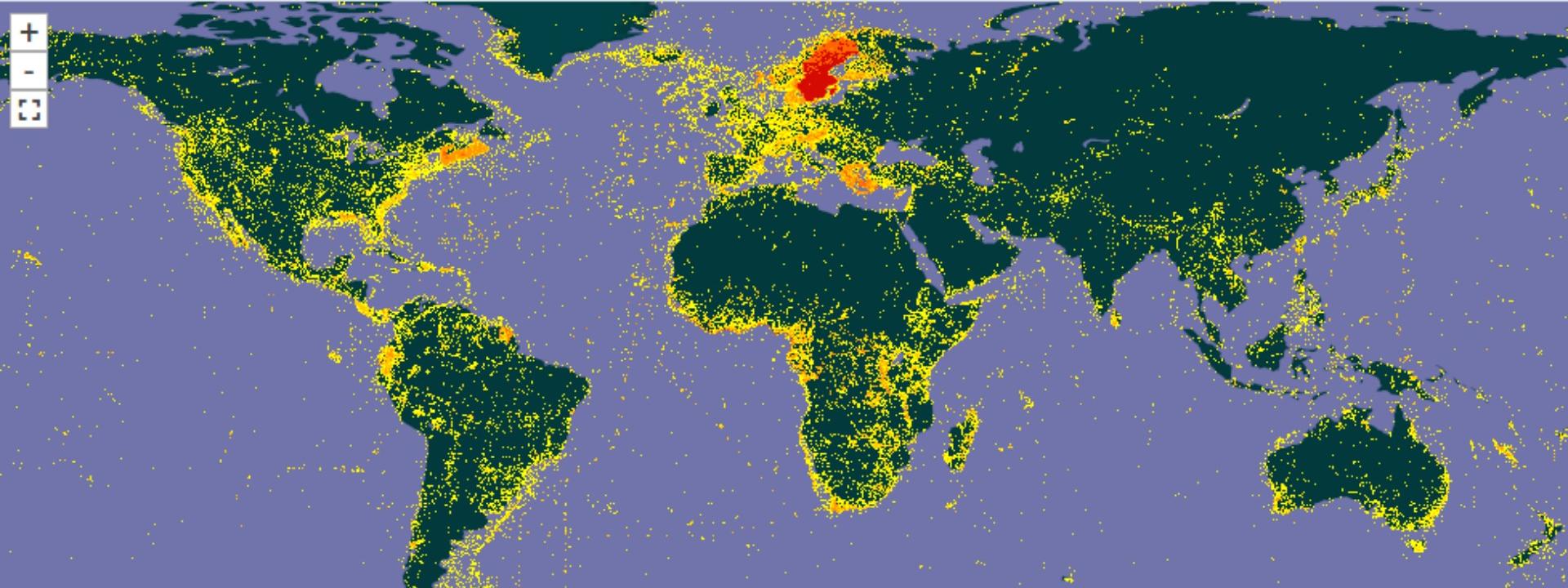
Published datasets

245

Countries and areas  
covered by data from  
Sweden

10

Publishers from Sweden



## **Short GBIF demo, finding specimen near Baikal**

- Short GBIF demo, finding specimen near Baikal from NRM



# Questions

Questions